Research Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.1215

Published online in Wiley Online Library

A Generalized Statistical Control Chart for Over- or Under-Dispersed Data

Kimberly F. Sellers[*†]

The Poisson distribution is a popular distribution used to describe count information, from which control charts involving count data have been established. Several works recognize the need for a generalized control chart to allow for data over-dispersion; however, analogous arguments can also be made to account for potential under-dispersion. The Conway–Maxwell–Poisson (COM-Poisson) distribution is a general count distribution that relaxes the equi-dispersion assumption of the Poisson distribution, and in fact encompasses the special cases of the Poisson, geometric, and Bernoulli distributions. Accordingly, a flexible control chart is developed that encompasses the classical Shewart charts based on the Poisson, Bernoulli (or binomial), and geometric (or negative binomial) distributions.

Keywords: attribute control chart; Conway–Maxwell–Poisson distribution; count data; Poisson distribution; Shewhart control chart

1. Introduction

The Poisson distribution has served as the benchmark for establishing much of the classical control chart theory, e.g. the Shewhart *c*- and *u*-charts. Several works (e.g. References[1–4]), however, express concern over the use of such a structure, recognizing that the underlying equi-dispersion assumption is limiting. Focusing on data sets that illustrate over-dispersion, the increased variation results in numerous data values being falsely detected as out-of-control when, in fact, such data are false positives. Sheaffer and Leavenworth[5], and Kaminsky *et al.* [6] consider the negative binomial distribution as an alternative to the Poisson control chart; in particular, Kaminsky *et al.*[6] directly develop the total number of events distribution (and the average number of events chart) via the geometric distribution.

Another framework, however, is also of concern; data can also display under-dispersion (i.e. where the variation is smaller than the mean). Under such circumstances, the limit bounds determined with a Poisson assumption would be too broad, thus implying that most data points are in-control because they fall within the associated bounds. Such data, however, could be false negatives (i.e. data values that are falsely interpreted as in-control when they are actually out-of-control) in that the decreased variation has not been assessed. Accordingly, under such circumstances, potential out-of-control states may (for example) require a longer study period to be detected. Famoye[7] introduces a statistical control chart based on a shifted generalized Poisson distribution of Consul and Jain[8] that can address data over- or under-dispersion. This model, however, gets truncated under certain conditions regarding the dispersion parameter and thus is not a true probability model[7].

Shmueli *et al.*[9] revived a flexible probability distribution called the Conway–Maxwell–Poisson (COM-Poisson) distribution that can broadly model count data that is either over- or under-dispersed. This distribution encompasses as special cases the Poisson, geometric, and Bernoulli distributions; Section 2 introduces this distribution in more detail. This distribution motivates a flexible control chart to describe the total number and average number of events, respectively. Section 3 describes the *cmpc* and *cmpu* control charts that generalize the classical Shewhart charts (namely, the *c*-chart and *u*-chart, the *np*-chart and *p*-chart, and the *g*-chart and *h*-chart described in Kaminsky *et al.*[6]) and serve as a bridge across the respective special cases. Section 4 provides simulation studies to demonstrate the model flexibility and performance, while Section 5 illustrates the flexibility of this model on various examples. Finally, Section 6 concludes the manuscript with a discussion.

Department of Mathematics and Statistics, Georgetown University, Washington, DC, U.S.A.
[*]Correspondence to: Kimberly F. Sellers, Department of Mathematics and Statistics, Georgetown University, Washington, DC, U.S.A.
[†]E-mail: kfs7@georgetown.edu

2. The COM-Poisson distribution

The COM-Poisson distribution (introduced by Conway and Maxwell[10], and revisited by Shmueli *et al.*[9]) is a viable count distribution that generalizes the Poisson distribution in light of associated data dispersion.

2.1. Distributional properties

The COM-Poisson probability mass function (pmf) takes the form

$$P(Y=y|\lambda,\nu)=\frac{\lambda^y}{(y!)^\nu Z(\lambda,\nu)},\quad y=0,1,2,\ldots, \tag{1}$$

for a random variable Y, where $\lambda=E(Y^\nu)$, ν is an associated dispersion parameter, and $Z(\lambda,\nu)=\sum_{s=0}^{\infty}\lambda^s/(s!)^\nu$ is a normalizing constant. The COM-Poisson distribution includes three well-known distributions as special cases: Poisson ($\nu=1$), geometric ($\nu=0$, $\lambda<1$), and Bernoulli ($\nu\to\infty$ with probability $\lambda/(1+\lambda)$). Accordingly, $\nu>1$ addresses the case of data under-dispersion, whereas $\nu<1$ implies over-dispersion.

In Shmueli *et al.*[9], the moments are given in the form

$$E(Y^{r+1})=\begin{cases}\lambda[E(Y+1)]^{1-\nu}, & r=0,\\ \lambda\dfrac{\partial}{\partial\lambda}E(Y^r)+E(Y)E(Y^r), & r>0,\end{cases} \tag{2}$$

and the expected value and variance are

$$E(Y)=\lambda\frac{\partial\log Z(\lambda,\nu)}{\partial\lambda}\approx\lambda^{1/\nu}-\frac{\nu-1}{2\nu}\ \text{and} \tag{3}$$

$$\mathrm{Var}(Y)=\frac{\partial E(Y)}{\partial\log\lambda}\approx\frac{1}{\nu}\lambda^{1/\nu}, \tag{4}$$

where the approximations particularly hold for $\lambda>10^\nu$[9, 11, 12]. The associated moment generating function of Y is $M_Y(t)=E(e^{Yt})=Z(\lambda e^t,\nu)/Z(\lambda,\nu)$, and its probability generating function is $E(t^Y)=Z(\lambda t,\nu)/Z(\lambda,\nu)$. Using the moment generating function, we see that for $\mathscr{Y}=\sum_{i=1}^n Y_i$ where $Y_i\sim$ COM-Poisson(λ,ν) are independent and identically distributed (i.i.d.) random variables ($i=1,\ldots,n$), the associated moment generating function is $M_{\mathscr{Y}}(t)=(Z(\lambda e^t,\nu)/Z(\lambda,\nu))^n$. This result is important because it generalizes the distributional results regarding the sum of i.i.d. Bernoulli, Poisson, and geometric random variables, respectively, thus forming a continuous bridge between the binomial, Poisson, and negative binomial distributions.

For this work, we are actually interested in the location-shifted form of the COM-Poisson distribution. Let $W=Y+a$ be a location-shifted COM-Poisson random variable, where Y has the pmf in Equation (1). Then W has pmf

$$P(W=w|\lambda,\nu)=\frac{\lambda^{w-a}}{[(w-a)!]^\nu Z(\lambda,\nu)},\quad w=a,a+1,a+2,\ldots. \tag{5}$$

Accordingly, the special cases are now shifted forms of their respective distributions, i.e. $\nu=1$ is a shifted Poisson(λ) distribution, $\nu=0$ is a shifted geometric (with $p=1-\lambda$) distribution as described by Kaminsky *et al.*[6], and $\nu\to\infty$ results in a shifted Bernoulli ($p=\lambda/(1+\lambda)$) distribution, i.e. W takes the value a [$a+1$] with probability $1/(1+\lambda)[\lambda/(1+\lambda)]$. Further, our mean is now $E(W)=\lambda\partial\log Z(\lambda,\nu)/\partial\lambda+a\approx\lambda^{1/\nu}+a-(\nu-1)/(2\nu)$, and the variance is as described in Equation (4).

2.2. Parameter estimation

We estimate the parameters λ, ν, and a via maximum likelihood estimation (MLE). Analogous to that described by Shmueli *et al.*[9], we maximize the likelihood function,

$$L(\lambda,\nu,a|\boldsymbol{w})=\prod_{i=1}^n\frac{\lambda^{w_i-a}}{[(w_i-a)!]^\nu Z(\lambda,\nu)}I_a(\boldsymbol{w}),$$

where $I_a(\boldsymbol{w})$ denotes the indicator function that attains the value one when all random variables are greater than or equal to a, and zero otherwise; accordingly, $\hat{a}=W_{(1)}$, the minimum order statistic. Meanwhile, to determine $\hat{\lambda}$ and $\hat{\nu}$, we optimize the likelihood function directly via `nlminb` in *R*.

3. A flexible control chart

Assume that a process generates events according to a shifted COM-Poisson (λ,ν) distribution. Letting X be the number of events per process unit, the probability distribution for X is given in Equation (5). Let $X_1,X_2,\ldots,X_n$ be a random sample of size n from the

Table I. Control limits for total number of events, and average number of events charts, respectively

	Total number of events chart	Average number of events chart
Centerline (CL)	$n\left(\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a\right)$	$\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a$
Upper control limit (UCL)	$n\left(\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a\right)+k\sqrt{n\left(\frac{\partial E(X)}{\partial \log\lambda}\right)}$	$\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a+k\sqrt{\frac{1}{n}\left(\frac{\partial E(X)}{\partial \log\lambda}\right)}$
Lower control limit (LCL)	$n\left(\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a\right)-k\sqrt{n\left(\frac{\partial E(X)}{\partial \log\lambda}\right)}$	$\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a-k\sqrt{\frac{1}{n}\left(\frac{\partial E(X)}{\partial \log\lambda}\right)}$

Table II. True model parameters versus model estimates (and associated goodness-of-fit p-values provided in parentheses) for various assumed distributions

	Estimated parameter			
True distribution	Poisson	Geometric	COM-Poisson	
Poisson($\lambda=10$)	$\lambda=9.986$	$p=0.091$	$\lambda=10.244$	$v=1.011$
	(0.9436)	(0.0000)	(0.8904)	
Geometric($p=0.2$)	$\lambda=3.862$	$p=0.206$	$\lambda=0.794$	$v=0.000$
	(0.0000)	(0.6220)	(0.6220)	
COM-Poisson($\lambda=10$, $v=5$)	$\lambda=1.184$	$p=0.458$	$\lambda=12.223$	$v=5.267$
	(0.0000)	(0.0000)	(0.4284)	
COM-Poisson($\lambda=3$, $v=0.5$)	$\lambda=9.510$	$p=0.095$	$\lambda=3.157$	$v=0.522$
	(0.0000)	(0.0000)	(0.6604)	

process, and consider $T=\sum_{i=1}^{n}X_i$ and $\bar{X}=T/n$, i.e. the total number of events, and the average number of events, respectively. Then,

$$E(T)=n\left(\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a\right),\quad \mathrm{Var}(T)=n\left(\frac{\partial E(X)}{\partial \log\lambda}\right),$$

$$E(\bar{X})=\lambda\frac{\partial \log Z(\lambda,v)}{\partial\lambda}+a,\quad \mathrm{Var}(\bar{X})=\frac{1}{n}\left(\frac{\partial E(X)}{\partial \log\lambda}\right).$$

Note that the approximations provided in Equations (3) and (4) may also be used under satisfying conditions or we can compute these terms exactly by using the `compoisson` package in *R*. Thus we can use the above, along with the parameter estimates determined as described in Section 2.2, to determine the centerline and $k\sigma$ control limits; see Table I.

Given the general form of the COM-Poisson distribution, the associated generalized control limits precisely equal the control limits for the special cases of *c*-chart and *u*-chart derived from the Poisson distribution (for $v=1$), and the *g*-chart and *h*-chart derived from the geometric distribution by Kaminsky *et al.*[6] (for $v=0$ and $\lambda<1$), and the *np*- and *p*-charts obtained from the Bernoulli distribution (as $v\to\infty$ with success probability $\lambda/(\lambda+1)$).

4. Simulation studies

The strength of the COM-Poisson distribution is its flexibility and ability to generalize the Poisson, geometric, and Bernoulli distributions. This is very helpful when the underlying count distribution is unknown. Accordingly, the COM-Poisson control chart theory presented here generalizes the *c*- and *u*-charts developed via the Poisson assumption, the *g*- and *h*-charts derived from the geometric distribution, and the *np*- and *p*-charts from the Bernoulli distribution, respectively. Thus, in-control data can be modeled well via the COM-Poisson control chart for good performance.

Below, we use the *c*-chart, *g*-chart, and COM-Poisson chart (*cmpc*-chart) to model data generated by several models, including the Poisson, geometric, and COM-Poisson distributions with some level of dispersion. We assess goodness of fit of the model on the sample data via the χ^2 test. Table II provides the model estimates obtained after generating data samples of size 500 from various models and then estimating them assuming particular distributional structures.

From Table II, we see that the Poisson and geometric models are only reasonably fit for the cases where the sample data were generated under that distributional assumption. Meanwhile, the COM-Poisson distribution provides a good fit irrespective of the distributional assumption. In fact, for the COM-Poisson simulations that demonstrate a different level of under- or over-dispersion, respectively, we see that the COM-Poisson model again is the only model able to reasonably fit this distribution.

Using the estimates provided in Table II, for simplicity, we determine the centerline and control limits by finding the respective mean and standard deviation for each case, thus computing the limits by the 'Mean $\pm k*$Std.Dev.' rule. For the Poisson and geometric distributions, the mean is easily determined; meanwhile, the COM-Poisson summary statistics are easily computed using the `compoisson` package in *R*. Alternatively, one can determine the upper control limit via a specified Type I error, α, so that the probability of being beyond the upper control limit (given the estimated parameters) equals α. In either case, in using historical in-control data to determine the control limits, we see that the COM-Poisson distribution can reasonably model the sample data demonstrating over- or under-dispersion, thus allowing for better flexibility when working with real data.

5. Examples

We consider three data sets containing various levels of dispersion to illustrate the flexibility of the COM-Poisson chart, and its ability to produce bounds comparable to those established from the classical statistical control theory. All computations were done using the statistical software, *R* (version 2.11.1), and here, we assume $k=3$.

5.1. *An example in under-dispersion*: *Location of chips on a wafer*

The National Institute for Standards and Technology provide the following example of a *p* chart on their web site (http://www.itl.nist.gov/div898/handbook/pmc/section3/pmc332.htm); interested persons are referred to the web site for the associated data. Thirty wafers are studied, and the number of defective chips (out of 50) on each wafer is measured. The data are used to create a *p* chart modeling the fraction of defective chips on the wafers. Here, we illustrate two results: (1) the COM-Poisson chart produces equivalent results to those obtained directly from the classical *p* chart procedure, and (2) the impact of incorrectly applying a *u*-chart to an under-dispersed data set.

In this example, we consider each chip independently, assessing whether the chip is defective or not. Accordingly, we first restructure the data so that it is composed of two columns, each of size $30\times 50=1500$: one tracking the sample or wafer number (1–30), and the other noting whether or not each of the 50 chips contained within that wafer is defective (1) or not (0). After restructuring, we determine the COM-Poisson MLEs with no shift (i.e. $\hat{a}=0$). Because we know that this structure is modeled via a Bernoulli distribution, we are reminded that the COM-Poisson distribution approaches a Bernoulli distribution as $\nu\to\infty$. Here, the optimization code converges at $\hat{\nu}\approx 30.17$ (with $\hat{\lambda}=0.3010$ i.e. $\hat{p}=\hat{\lambda}/(\hat{\lambda}+1)=0.2313$) to obtain identical results to those from the *p*-chart obtained via classical methods; see Figure 1. Wafers 15 and 23 are thereby deemed out-of-control because they fall beyond the upper bound determined either via the *p*-chart or *cmpu*-chart.

Meanwhile, *u*-chart [i.e. Poisson ($\hat{\lambda}=0.2313$)] computations produce a centerline equal to that obtained via the *cmpu*- and *p*-charts, yet obtains upper and lower limit bounds that are wider than its counterparts. The upper and lower bounds are approximately 0.4948 and 0, respectively, where the lower bound is Winsorized to zero because the true lower bound computation results in a negative value. This results in Wafers 15 and 23 being viewed as in-control, while Wafer 23 may still warrant an investigation if using an upper warning limit at two standard deviations above from the centerline.

5.2. *An example in over-dispersion*: *A simulated geometric data example*

Kaminsky *et al.*[6] developed a Shewhart chart based on the (shifted) geometric distribution and illustrated the repurcussions of misapplying a *c*-chart to such data. Coining the terms *g*- and *h*-charts to reference the total number and average number of events, respectively, they show that a *c*-chart misidentifies sample points as being out-of-control where the corresponding *g*-chart recognizes the inherent over-dispersion in the data and modifies the control limits accordingly to accommodate the added variation. As a result, all points fall within the *g*-chart limits. Famoye[7] develops a control chart using a shifted generalized

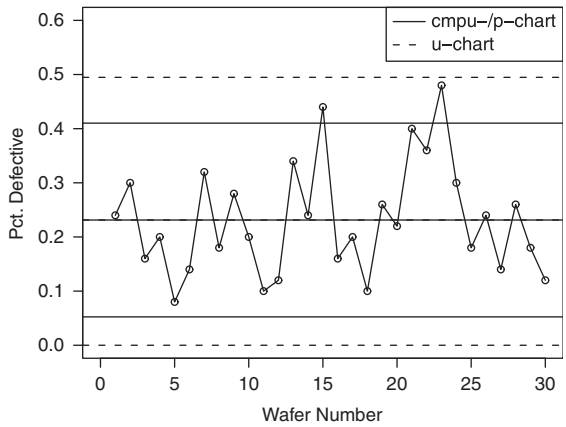

Figure 1. Control chart comparisons: *u*-chart, *p*-chart, and *cmpu*-chart. The *p*-chart and *cmpu*-chart bounds (with $\hat{p}=\hat{\lambda}/(\hat{\lambda}+1)=0.3010/1.3010=0.2313$; $\hat{\nu}=30.1668$) are identical. The *u*-chart [obtained via the Poisson(0.2313) distribution] produces wider bounds because the data are under-dispersed

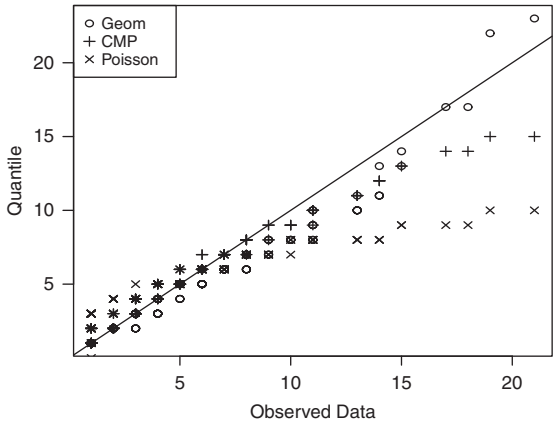

Figure 2. QQ-plot comparing original data from Kaminsky *et al.*[6] with the geometric($p=0.2$), Poisson($\hat{\lambda}=5.0667$), and COM-Poisson($\hat{\lambda}=0.8211$, $\hat{\nu}=0.0118$) distributions

Table III. Control limits for simulated geometric data determined via various Shewhart charts (where $k\sigma=3\sigma$): *c*-chart, *g*-chart, *gc*-chart, *cmpc*-chart

	True value	*c*-chart	*g*-chart	*gc*-chart	*cmpc*-chart
Upper limit	55.00	40.43	55.73	56.63	55.01
Centerline	25.00	25.33	25.33	25.33	25.33
Lower limit	5.00	10.23	5.00	5.00	5.00

Poisson distribution to likewise model this data set, with interesting results. The shifted generalized Poisson distribution proved to fit this data set well, and the corresponding *gc*-chart produced limits comparable to those determined by Kaminsky *et al.*[6]. Here, we likewise illustrate the effectiveness of the COM-Poisson distribution in modeling geometric data, as expected given that the geometric distribution is a special case of the COM-Poisson distribution ($\nu=0$).

Kaminsky *et al.*[6] simulated data from a shifted geometric distribution (with success probability $p=0.20$ and shift $a=1$). Using our optimization scheme on their data, we find the COM-Poisson MLEs equal $\hat{\lambda}=0.8211$, $\hat{\nu}=0.0118$, and $\hat{a}=1$. The estimate $\hat{\nu}$ shows that the data are likewise recognized as being approximately geometric in nature. In fact, when $\nu=0$, the shifted COM-Poisson distribution reduces to a shifted geometric distribution with $p=1-\lambda$; accordingly, we can see that the COM-Poisson MLEs reasonably approximate the true distribution; see Figure 2. The *cmpc*-chart likewise produces control limits comparable to the *g*-chart of Kaminsky *et al.*[6] and *gc*-chart of Famoye[7]. In fact, the centerline and lower limits are equal for all three methods, while the upper limits are approximately equal. In particular, the lower limit is Winsorized to five in all cases because the corresponding computed lower bound is less than 5. Taking into account the true simulation values, however, reveals that the true centerline is 25 with lower and upper control limits 5 and 55, respectively; see Table III. Accordingly, the *cmpc*-chart produces values closest to the true limits. Figure 3 shows the statistical control charts determined by these various methods.

5.3. To c or not to c?

Montgomery[13] provides an example utilizing the *c*-chart where non-conformities are observed in 26 samples of circuit boards. Using the classical *c*-chart method, we find that the MLE for the (assumed) underlying Poisson distribution is $\hat{\lambda}=19.8462$ (i.e. the Poisson mean and standard deviation are 19.8462 and 4.4549, respectively), making this the centerline value, and the associated control limits are 6.4814 and 33.2109, respectively; see Figure 4. Accordingly, two points fall outside of the control limits (namely Observations 6 and 20). Montgomery[13], in turn, provides insight into the causes of the respective out-of-control samples (which are subsequently corrected) and argues for their exclusion in order to obtain revised trial control limits.

When using the original data and applying the optimization scheme to determine the COM-Poisson MLEs, however, we find that $\hat{\lambda}=2.8711$ and $\hat{\nu}=0.3652$, i.e. the data set is over-dispersed. Figure 5 provides the quantile plots comparing the sample data with the Poisson($\hat{\lambda}=19.462$) and COM-Poisson($\hat{\lambda}=2.8711$, $\hat{\nu}=0.3652$) distributions, respectively. While the Poisson distribution works to capture the large percentage of sample data containing between 15 and 25 non-conformities, its constraining equidispersion assumption does not allow it to better capture either tail in the sample distribution. Meanwhile, the COM-Poisson distribution seems to better model the full distribution of non-conformities in the sample data. The necessary constraint to approximate the COM-Poisson mean and variance is satisfied (namely $\hat{\lambda}>10^{\hat{\nu}}$); hence, we use the approximations provided in Equations (3) and (4) to determine the associated control limits; the mean and standard deviation are 18.8270 and 7.0123, respectively. Therefore, the *cmpc*-chart centerline is 18.8270 with an upper control bound of 39.8640. Meanwhile, the lower bound is Winsorized to zero because the computed lower bound is negative; see Figure 4.

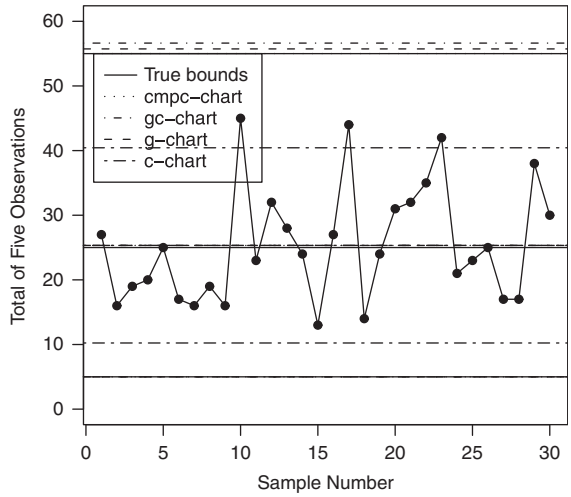

Figure 3. Control chart bounds derived from various methods: *c*-chart, *g*-chart, *gc*-chart, and *cmpc*-chart. *cmpc*-chart bounds are nearly equivalent with (and thus hidden under) true bounds. The lower bounds attained by the true model, *g*-, *gc*- and *cmpc*-charts are Winsorized to zero

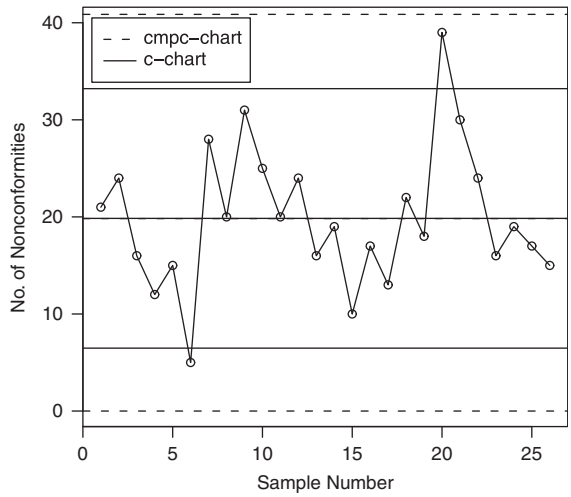

Figure 4. Control chart bounds derived via *c*-chart and *cmpc*-chart methods. The *cmpc*-chart lower bound is Winsorized to zero

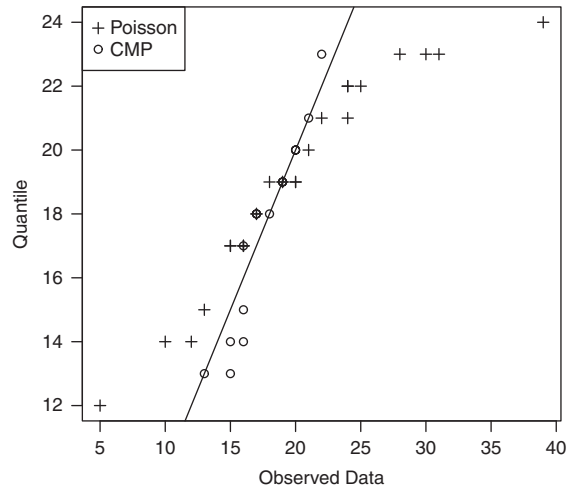

Figure 5. Distribution of circuit board non-conformities. QQ-plot compares original data with Poisson($\hat{\lambda}=19.8462$) and COM-Poisson($\hat{\lambda}=2.8711$, $\hat{\nu}=0.3652$) distributions

An interesting byproduct of this exercise is the wider bounds associated with the COM-Poisson procedure; the centerlines determined from both models are approximately equal (19.8462 for the Poisson centerline versus 19.8270 for the COM-Poisson centerline). While the Poisson assumption allows one to capture the out-of-control samples and rectify the apparent quality issues, the COM-Poisson procedure believes these samples to be in-control. At the same time, however, Sample 20 would presumably still draw attention because it is so close to the upper limit and is beyond an upper warning limit at two standard deviations above the centerline (the upper warning bound is 33.8517). Further, this example demonstrates the importance of historical assumptions. If the examiners have reason to believe (as was the case here) that the data presumably stem from a Poisson distribution, then the COM-Poisson approach would produce estimates, $\hat{\lambda}=19.8462$ and $\hat{\nu}=1$, i.e. equivalent to the Poisson estimate. Accordingly, the underlying assumption produces *cmpc*-chart and *c*-chart bounds that are equivalent. Thus, the results would likewise be equivalent in either case. If no such distributional assumption exists, however, and we have to assume that this data set is in-control, then we find the samples perceived to be out-of-control under the Poisson assumption to be invalid.

6. Discussion

The COM-Poisson distribution is a flexible distribution that generalizes the Poisson, geometric, and Bernoulli distributions. Accordingly, the control charts associated with these distributions likewise are special cases of the COM-Poisson control chart described here. As witnessed by the examples in Section 5, the amount of dispersion influences the size of the lower and upper bounds, thus impacting one's inference regarding the status of any sample points. If the data are under-dispersed, *c*- or *u*-charts risk falsely identifying points as in-control, thereby extending the time until the process is recognized as out-of-control. Meanwhile, if a data set is over-dispersed, *c*- and *u*-charts may prematurely denote samples as out-of-control when, in reality, added variation naturally exists in the data.

What does this mean for the analyst? The problem of Type I versus Type II errors is an age-old problem. One must assess the impact of making one of these errors in determining the choice of chart construction. To circumvent this problem, one can use historical in-control data to determine the control limits. Under such circumstances, the beauty of this approach lies in the COM-Poisson's ability to provide an extra layer of flexibility and generalizability that does not exist with the other (i.e. *c*-, *u*-, *np*-, *p*-, *g*-, and *h*-) charts.

In this work, we determine the limits based on a three-sigma rule. Alternatively, one can determine the upper control limit via a specified Type I error, α, so that the probability of being beyond the upper control limit (given the estimated parameters) equals α.

References

1. Spiegelhalter D. Handling over-dispersion of performance indicators. *Quality and Safety in Health Care* 2005; **14**:347–351.
2. Mohammed M, Laney D. Overdispersion in health care performance data: Laney's approach. *Quality and Safety in Health Care* 2006; **15**(5):383–384.
3. Albers W. Control charts for health care monitoring under overdispersion. *Metrika* 2009; DOI: 10.1007/s00184-009-0290-z.
4. Chen N, Zhou S, Chang T-S, Huang H. Attribute control charts using generalized zero-inflated Poisson distribution. *Quality and Reliability Engineering International* 2008; **24**(7):793–806.
5. Sheaffer R, Leavenworth R. The negative binomial model for counts in units of varying size. *Journal of Quality Technology* 1976; **8**(3):158–163.
6. Kaminsky FC, Benneyan JC, Davis RD, Burke RJ. Statistical control charts based on a geometric distribution. *Journal of Quality Technology* 1992; **24**(2):63–69.
7. Famoye F. Statistical control charts for shifted generalized Poisson distribution. *Statistical Methods and Applications* 2007; **3**(3):339–354.
8. Consul P, Jain G. A generalization of the Poisson distribution. *Technometrics* 1973; **15**:791–799.
9. Shmueli G, Minka TP, Kadane JB, Borle S, Boatwright P. A useful distribution for fitting discrete data: Revival of the Conway–Maxwell–Poisson distribution. *Applied Statistics* 2005; **54**:127–142.
10. Conway RW, Maxwell WL. A queuing model with state dependent service rates. *Journal of Industrial Engineering* 1962; **12**:132–136.
11. Guikema SD, Coffelt JP. A flexible count data regression model for risk analysis. *Risk Analysis* 2008; **28**(1):213–223.
12. Minka T, Shmueli G, Kadane J, Borle S, Boatwright P. Computing with the COM-Poisson distribution. *Technical Report 776*, Department of Statistics, Carnegie Mellon University, 2003.
13. Montgomery D. *Introduction to Statistical Quality Control* (6th edn). Wiley: New York, 2008.

Authors' biographies

Kimberly F. Sellers received her BS (1994) and MA (1998) in Mathematics from the University of Maryland College Park, and her PhD. (2001) in Statistics from The George Washington University in Washington, DC. She is currently an Assistant Professor of Statistics in the Department of Mathematics and Statistics at Georgetown University, also in Washington, DC. Her research involves generalized statistical methods involving count data in the light of data dispersion.